

Frauke Stanke · Tim Becker · Harry Cuppens
Vinod Kumar · Jean-Jacques Cassiman
Silke Jansen · Dragica Radojkovic
Benny Siebert · Jennifer Yarden · David W. Ussery
Thomas F. Wienker · Burkhard Tümmler

The TNF α receptor TNFRSF1A and genes encoding the amiloride-sensitive sodium channel ENaC as modulators in cystic fibrosis

Received: 8 September 2005 / Accepted: 20 December 2005
© Springer-Verlag 2006

Abstract The *CFTR* mutations in cystic fibrosis (CF) lead to ion transport anomalies which predispose to chronic infection and inflammation of CF airways as the major determinants for morbidity and mortality in CF. Discordant clinical phenotypes of siblings with identical *CFTR* mutations and the large variability of clinical manifestations of patients who are homozygous for the most common mutation F508del suggest that both environment and genes other than *CFTR* contribute substantially to CF disease. The prime candidates for genetic modifiers in CF are elements of host defence such as the TNF α receptor and of ion transport such as the amiloride-sensitive epithelial sodium channel ENaC, both of which are encoded side by side on 12p13 (*TNFRSF1A*, *SCNN1A*) and 16p12 (*SCNN1B*, *SCNN1G*). Thirty-seven families with F508del-*CFTR* homozygous siblings exhibiting extreme clinical phenotypes that had been

selected from the 467 pairs of the European CF Twin and Sibling Study were genotyped at 12p13 and 16p12 markers. The ENaC was identified as a modulator of CF by transmission disequilibrium at *SCNN1G* and association with CF phenotype intrapair discordance at *SCNN1B*. Family-based and case-control analyses and sequencing of *SCNN1A* and *TNFRSF1A* uncovered an association of the *TNFRSF1A* intron 1 haplotype with disease severity. Carriers of risk haplotypes were underrepresented suggesting a strong impact of both loci on survival. The finding that *TNFRSF1A*, *SCNN1B* and *SCNN1G* are clinically relevant modulators of CF disease supports current concepts that the depletion of airway surface liquid and inadequate host inflammatory responses trigger pulmonary disease in CF.

Keywords Cystic fibrosis · Modulating gene · Amiloride-sensitive epithelial sodium channel · TNF α receptor · Association study · Survivor effect

Electronic Supplementary Material Supplementary material is available for this article at <http://dx.doi.org/10.1007/s00439-006-0140-2> and is accessible for authorized users.

F. Stanke (✉) · V. Kumar · S. Jansen · B. Siebert · B. Tümmler
Department of Pediatrics, OE6711, Hannover Medical School,
Carl-Neuberg-Str. 1, 30625 Hannover, Germany
E-mail: mekus.frauke@mh-hannover.de
Tel.: +49-511-5326722
Fax: +49-511-5326723

T. Becker · T. F. Wienker
Institute for Medical Biometry, Informatics and Epidemiology,
University of Bonn, Bonn, Germany

H. Cuppens · J.-J. Cassiman · D. Radojkovic · J. Yarden
Department of Human Genetics, KU Leuven,
Leuven, Belgium

D. Radojkovic
Institute of Molecular Genetics and Genetic Engineering,
Belgrade, Serbia and Montenegro

D. W. Ussery
Center for Biological Sequence Analysis,
The Technical University of Denmark,
Lyngby, Denmark

Introduction

Cystic fibrosis (CF) is the most common severe monogenic disease inherited in an autosomal recessive fashion within the Caucasian population with an incidence of 1 in 2,500 births (Welsh et al. 1995). The disease is caused by molecular lesions in the cystic fibrosis transmembrane conductance regulator gene *CFTR*, encoding an apical chloride channel that is expressed in epithelial cells of exocrine glands (Welsh et al. 1995). The multi-organ CF disease presents as a generalized exocrinopathy whereby the major affected organs are the gastrointestinal tract and the lung (Welsh et al. 1995). Continuously improving treatment regimens have steadily increased the patients' survival and quality of life over the last decades (Littlewood 2000; Goss and Rosenfeld 2004).

While more than 1,300 sequence variants of the *CFTR* gene are known (The CF mutation database 2004), half of the patients of Caucasian descent are

homozygous for the most common CF mutation F508del-*CFTR* (European Working Group on Cystic Fibrosis Genetics 1990). Course and prognosis of CF disease varies considerably among these individuals (Johannsen et al. 1991; Kerem et al. 1990) and in affected patient pairs (Santis et al. 1990; Mekus et al. 2000). This variability of CF disease has sparked numerous hypotheses on the impact of modulating genes which have been investigated by interrogating known variants in selected candidate genes (Cutting 2005).

Combining two quantitative clinical parameters validated to reflect nutritional and pulmonary status in CF, we have previously described an algorithm to select informative pairs of F508del homozygous siblings based on their overall disease severity from the *European CF Twin and Sibling Study* (Mekus et al. 2000, 2003). Here, we have employed the phenotypic contrast between these selected F508del-*CFTR* homozygous sib pairs exhibiting concordant mildly, concordant severely and discordant phenotypes to ask whether genes encoding the subunits of the amiloride-sensitive epithelial sodium channel ENaC (Baens et al. 1995; Voilley et al. 1995) modulate CF.

The targeted candidate gene ENaC is a plausible choice for a CF modulator because first, an abnormal chloride transport constitutes the *CFTR*-mediated basic defect (Welsh et al. 1995) and ENaC transports the sodium counterion, second, interactions between the two ion channels ENaC and *CFTR* have been shown at the molecular level (Ji et al. 2000) and third, mutations within the ENaC subunit genes are known to cause pseudohypoaldosteronism type 1 (OMIM 264350; Chang et al. 1996; Strautnieks et al. 1996), a disease which shares some symptoms with CF such as elevated sweat electrolytes and *Pseudomonas aeruginosa* infections in the lung (Hanukoglu et al. 1994; Marthinsen et al. 1998).

SCNN1B and *SCNN1G* encoding the β - and γ -subunits of ENaC are located within 200 kb on 16p12 (Voilley et al. 1995). The gene encoding the α -subunit of ENaC, *SCNN1A*, maps to 12p13 (Baens et al. 1995) and is the neighboring gene to *TNFRSF1A* (International Human Genome Sequencing Consortium 2003), encoding the 55 kDa receptor for the cytokine TNF α (Fuchs et al. 1992). *SCNN1A* and *TNFRSF1A* are separated by 5 kb of intergenic sequence. The close physical distance of these genes demands that *TNFRSF1A* is implemented in the analysis for CF modulators on 12p13. The 55 kDa receptor for TNF α mediates the pleiotropic actions of this proinflammatory cytokine on its target cells (Wajant et al. 2003), and thus *TNFRSF1A* constitutes a prominent candidate gene for the modulation of CF disease itself. In this context, it is interesting to note that CF is associated with a proinflammatory state which is mediated by the TNF α /NF κ B pathway (Eidelman et al. 2001; Tabary et al. 2001).

In order to address whether or not modulators of CF disease severity are encoded on 12p13 and 16p12 and to dissect the relative impact of each out of the two

neighboring candidate genes on both chromosomal locations, we have typed 37 families with F508del-*CFTR* homozygous siblings exhibiting extreme clinical phenotypes (Mekus et al. 2000) at 20 SNP and microsatellite markers.

Patients and methods

Recruitment of F508del-*CFTR* homozygous patients

The families enrolled for this study were selected from the set of CF patient pairs who participated in the *European CF Twin and Sibling Study* as described previously (Mekus et al. 2000). We chose to select extreme clinical phenotypes for analysis of modulators of CF disease severity based on their informativeness in mapping quantitative traits (Eaves and Meyer 1994; Risch and Zhang 1995; Dolan and Boomsma 1998). Briefly, we have selected dizygous F508del homozygous CF siblings exhibiting extreme clinical phenotypes as judged by a ranking algorithm that relies on two clinical parameters most sensitive for course and prognosis of CF, i.e. weight as % of predicted weight for height (wfh%, representative for the nutritional status) and CF population centiles for the forced expiratory volume in 1 s, expressed as % of predicted values (FEVPer, representative for the pulmonary status).

The pairs were selected from a group of 318 patient pairs, of whom 114 were homozygous for the most frequent CF disease-causing lesion F508del-*CFTR*. Clinical data for the siblings were obtained in 1995–1996 from 158 CF clinics from 14 European countries. Pairs were ranked by their disease severity and intrapair discordance to identify informative phenotypes wherein both siblings exhibit severe disease (concordant/severe disease; CON–), or both siblings exhibit mild disease (concordant/mild disease; CON+) or pairs wherein one sibling is mildly affected while the other is severely affected (discordant; DIS). Severely affected sibs CON– and mildly affected sibs CON+ differ significantly in both clinical parameters as utilized by the ranking algorithm (mean wfh%: 90.7 for CON–, 94.3 for DIS and 107.9 for CON+; mean FEVPer: 15.9 for CON–, 23.3 for DIS and 53.5 for CON+; see Mekus et al. 2000, 2003 for details). Discordant DIS sibpairs differed significantly from the concordant phenotypes (CON+ and CON–) in their intrapair-difference for both clinical parameters (mean intrapair difference for wfh%: 6.0 for CON–, 6.5 for CON+ and 19.9 for DIS; mean intrapair difference for FEVPer: 13.5 for CON–, 19.0 for CON+ and 47.9 for DIS). The selected patient pairs represent the extreme clinical phenotypes among the F508del homozygous population of the *European CF Twin and Sibling Study* whereby the selected severely and mildly affected phenotypes display phenotypes below the 25th centile and above the 75th centile, respectively.

The gender and age distribution among the mildly affected, severely affected and discordant siblings was comparable. The year of birth (median [inner quartiles, range]) was: 1982 [1977–1985, 1963–1990] for CON–; 1981 [1977–1986, 1962–1990] for CON+; 1977 [1971–1983, 1957–1990] for DIS.

The dizygous F508del homozygous patient pairs investigated within this study consist of 11 CON– families, 12 CON+ families and 14 DIS families. DNA of at least one parent was available for 32 of the 37 sib pairs (Table 1). Thirty-five of the 37 pairs were dizygous sibling pairs while two pairs were dizygous twins. Data from monozygous twins were not used in the analysis of genotype–phenotype associations.

Mendelian inheritance of the F508del mutation was verified by genotyping the affected F508del homozygous sibs and their F508del heterozygous parents by heteroduplex analysis (Rommens et al. 1990). Siblings shared their *CFTR* genetic background as no recombination event was observed at loci MetH and J3.11, two diallelic markers located in 1 Mb distance to the 3' and the 5' ends of the *CFTR* gene, respectively (Mekus et al. 2003). Non-paternity and consanguinity were excluded by genotyping of 16 polymorphic markers located on five autosomes other than chromosome 7. No biased overrepresentation of homozygous genotypes as an indicator of consanguinity was observed (Mekus et al. 2003).

The study was approved by the local medical ethics committees and informed consent was obtained from the participating patients and parents.

Genetic markers and haplotypes

High-molecular weight DNA was isolated from K-EDTA–blood by standard procedures (Gross-Bellard et al. 1973). To analyze the 0.5 Mb region including *SCNNIA* and *TNFRSF1A* on chromosome 12p13, seven SNPs and four microsatellites were typed (Fig. 1a, Table 4). The 16p12 region harboring *SCNNIB* and *SCNNIG* was interrogated with 6 SNPs and 3 microsatellite markers encompassing 0.8 Mb (Fig. 1b, Table 4). Genotyping was carried out using PCR-based methods as specified in the Appendix. Microsatellite markers were selected based on their proximity to the candidate genes *SCNNIA* (D12S889; Baens et al. 1995) and *SCNNIB* (betaENaCGT; Shimkets et al. 1994), or by their reported position on the annotated genomic

Table 1 No of families with F508del homozygous patient pairs selected for genotyping

| Phenotype | Sibs only | Sibs and one parent | Sibs and both parents | All families |
|-----------|-----------|---------------------|-----------------------|--------------|
| CON– | 2 | 1 | 8 | 11 |
| CON+ | – | 2 | 10 | 12 |
| DIS | 3 | 1 | 10 | 14 |

CON– concordant severely affected patient pair, CON+ concordant mildly affected patient pair, DIS discordant patient pair

contigs deposited within the NCBI database (D12S374 and D16S417), or by screening genomic sequences for repeat motifs (HS3ST2Sat, SC3 and SC4). All microsatellite markers were typed using PCR amplification with one biotinylated primer and high-resolution direct blotting electrophoresis with subsequent chemoluminescence detection as described elsewhere (Mekus et al. 1995). All SNPs but nt7 A/G within *SCNNIA*, the latter being identified by direct sequencing of *SCNNIA* exon 1 (unpublished data) were selected from the NCBI database based on their location within the regions of interest. For typing of SNPs rs3181301, rs740842, rs1800693, rs1800692, rs767455, rs5723 and rs152730, PCR-RFLP protocols were developed. The SNPs at nt7 of *SCNNIA*, rs2228576, rs5735, rs1004749, rs238547, rs250563 were typed by single nucleotide primer extension assay using the SNP detection kit (SnaPshot™—Applied Biosystems).

Alleles at all SNP loci denote the respective nucleotide. Alleles at microsatellite loci are named according to the number of repeat units with allele 10 at D12S889 corresponding to (GA)₁₀ and allele 13 at D12S889 corresponding to (GA)₁₃. Alleles at markers SC3, SC4 and betaENaCGT are given as arbitrary repeat units reflecting the relative size of detected PCR products. An invariant set of calibrated control samples was run on all gels to ascertain the unequivocal identification of microsatellite alleles.

Consistency of genotyping was checked for by the criterion of Mendelian inheritance within families. A set of five monozygous twins was typed at all microsatellites and all SNP loci evaluated by PCR-RFLP.

Haplotypes, integrating information of more than one marker, are described throughout this manuscript by specifying the loci considered and stating the alleles observed at these positions on the respective chromosomal background. For instance, the combination of allele C at rs5735 and allele C at rs5723 is expressed as two marker haplotype CC at rs5735–rs5723, thus describing the chromosomal segment between codon 158 (rs5735) and codon 649 (rs5723) of *SCNNIG*. Details on the location of SNP and microsatellite loci are given in the legend to Fig. 1 and in Table 4 of the supplement. The position and gene context of the first and the last marker of each haplotype is given in brackets, e.g. rs5735–rs5723 [codon 158 to codon 649 of *SCNNIG*], to identify the genomic segment described. The sequence of loci are given based on ascending positions in the contigs NT_009759 (Version April 2003) for 12p13 markers and NT_010393 (Version August 2004) for 16p12 markers whereby on NT009759, *FLJ10665*, *TNFRSF1A* and *SCNNIA* are represented on the complementary strand.

Data analysis

Three different hypothesis were tested on 12p13 and on 16p12. Firstly, we tested for association with the affection status as defined by the ascertainment strategy of

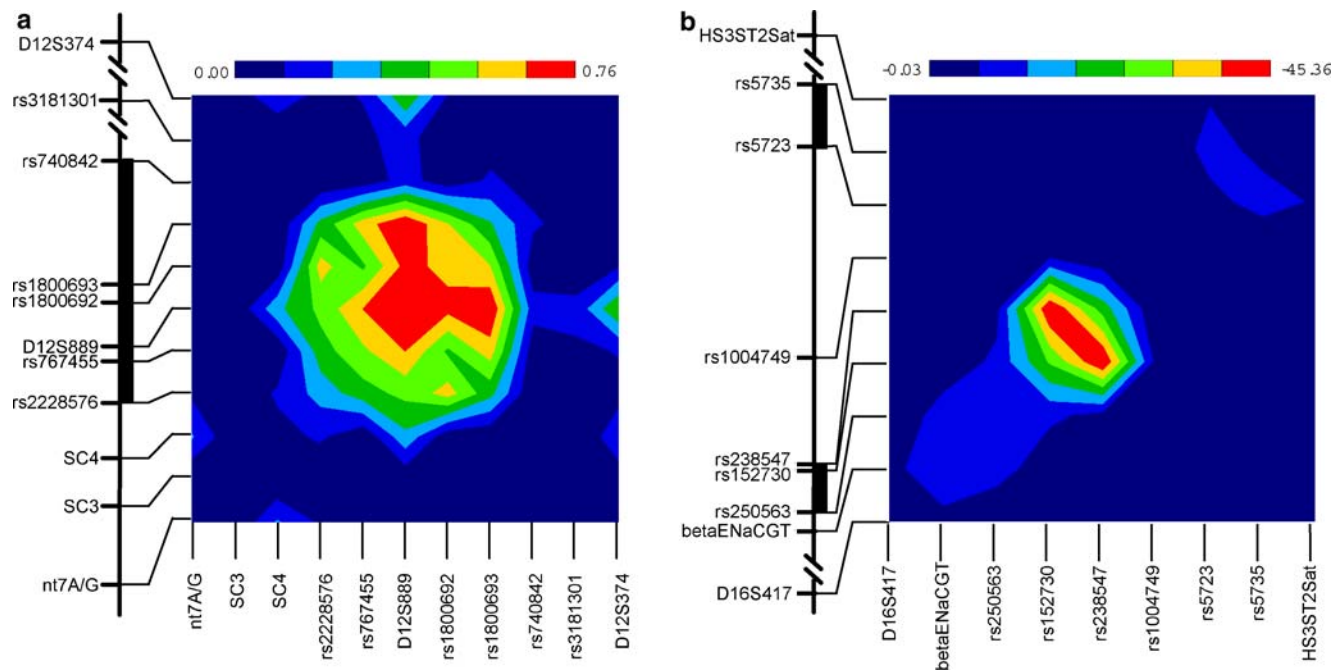


Fig. 1 Measure of linkage disequilibrium. Pairwise LD distribution for SNP and microsatellite markers analyzed on 12p13 (a) and 16p12 (b). *P* values for the Chi-square measure are visualized as indicated by the color code. Solid bars on the physical map designate the haplotype blocks. The physical map is drawn to scale for rs740842–nt7A/G on 12p13 showing a 62.3 kb segment and for rs5735–betaENaCGT on 16p12 showing a 187 kb segment. The physical map is not drawn to scale for markers D12S374, rs3181301, HS3ST2Sat and D16S417. For these markers, distances to neighboring markers on the physical map are: D12S374–rs3181301: 379 kb; rs3181301–rs740842: 76 kb; HS3ST2Sat–rs5735: 324 kb; betaENaCGT–D16S417: 390 kb. D12S374, HS3ST2Sat and D16S417 are intergenic markers while all other

SNPs or microsatellites are located within genes *CD9*, *FLJ10665*, *TNFRSF1A* and *SCNN1A* on 12p13 or *SCNN1G* and *SCNN1B* on 16p12. On 12p13, rs3181301 is in intron 7 of *CD9*; rs740842 corresponds to T35A in *FLJ10665*; rs1800693 is in intron 6, rs1800692 is in intron 5 and D12S889 is in intron 1 of *TNFRSF1A*; rs767455 corresponds to P12P in *TNFRSF1A*; rs2228576 corresponds to T663A in *SCNN1A*, SC4 is in intron 4 and SC3 is in intron 3 of *SCNN1A*, nt7 A/G is in the 5' UTR of *SCNN1A*. On 16p12, rs5735 corresponds to I158I and rs5723 corresponds to L649L in *SCNN1G*; rs1004749 is in the 5' UTR of *SCNN1B*; rs238547 corresponds to P3P in *SCNN1B*; rs152730 is in intron 1 of *SCNN1B*, rs250563 corresponds to F293F in *SCNN1B*, betaENaCGT is in intron 6 of *SCNN1B*

the CF siblings. For this purpose, transmissions/non-transmissions were considered in the joint sample of all nuclear families, irrespective of the grouping into different classes of severity. Since only affected sib-pairs for which both sibs survived at least until the age of 7 years when lung function measurements can reliably be performed were enrolled, preferential transmission of certain alleles (or haplotypes) yields evidence for a mechanism which modifies the prognosis for survival of CF patients in utero and early postnatal life. Secondly, we compared the sample of CON+ siblings to the sample of CON– siblings, employing the phenotypic contrast to identify CF modifiers. Thirdly, the pooled sample of CON+ and CON– siblings was compared to the group of DIS siblings. Differences in the frequency of the analyzed DNA variation point to a target for interacting genes elsewhere in the genome, which modify disease severity.

Pairwise marker linkage disequilibrium (LD) was judged both on the basis of Lewontin's *D'* measure (Lewontin 1964) as well as on the significance of the Chi-square measure. For the latter, the contingency table was calculated and its significance was approximated via

a Chi-square distribution with $(s-1)(t-1)$ degrees of freedom, where *s* is the number of alleles of marker 1 and *t* is the number of alleles of marker 2. It turned out that several markers showed little or no evidence for LD with any of the other markers. In particular, the microsatellite markers, except for D12S889, fell into this category. Therefore, we distinguished conceptually between a single-marker analysis and a multi-marker analysis. Multi-marker analyses were carried out only within clearly identifiable blocks of markers with elevated LD. On chromosome 12, one block was identified, while two different blocks could be determined on chromosome 16 (Fig. 1).

Nuclear families were analyzed with the Monte Carlo simulation based association test described by Knapp and Becker (2003), which can be viewed as an extension of the transmission–disequilibrium test (Spielman et al. 1993) to both nuclear families with more than one affected child and to haplotypes. Furthermore, we conducted tests of case-control type and compared marker/haplotype frequencies between two groups of sib pairs in order to test our second and third hypothesis. We used the Chi-square statistic for

contingency tables and determined significance via Monte Carlo simulations. In order to account for the dependence of the genotypes of the individuals within each sib pair, in each permutation replicate the affection status was simultaneously permuted or not permuted with equal probability for both sibs. Haplotype counts were obtained by using lists of likelihood-weighted haplotype explanations for each individual as described in Becker et al. (2005). Additionally, we carried out a genotype/diplotype inspired testing strategy. Each allele or haplotype of a given marker combination was evaluated both under a recessive and a dominant model. For each allele (haplotype) we considered a 2x2 contingency table. Cell counts were deduced from the weighted list of haplotypes and then Chi-square statistic was computed for each of the two models. For each marker combination, we applied the same Monte Carlo simulations procedure as described before to account for the dependence of the genotypes of the siblings. We simultaneously corrected for the number of different haplotypes and models at each marker combination by the use of the maximum statistic of all 2x2 Chi-square values.

The multiple testing issue induced by the consideration of different markers or marker combinations was addressed as follows: within each region, we applied a Monte Carlo simulation based strategy to obtain a correction for the number of single markers tested to obtain a common *P* value for the single marker analysis. On top of this, within each block of elevated LD we considered all marker combinations and tested them for association with the disease. The multiple testing issue was addressed as described in Becker and Knapp (2004b) and a common *P* value for each block was obtained.

All computations were carried out using the FAM-HAP software package (Becker and Knapp 2004a).

Results

Thirty-seven F508del-*CFTR* homozygous sib pairs and their families were genotyped at 12p13 and 16p12 markers. The 16p12 area encompasses *SCNNIG* and *SCNNIB* encoding the γ - and β -subunits, respectively, of the epithelial amiloride-sensitive sodium channel. The investigated area on 12p13 contains *TNFRSF1A*, encoding the 55 kDa subunit of the TNF α -receptor, and *SCNNIA*, encoding the α -subunit of the epithelial amiloride-sensitive sodium channel. For all of these four genes, an impact on CF disease severity can be reasonably assumed based on the biological function of the gene products. A family-based test comparing transmitted and non-transmitted alleles was applied to all nuclear families to detect genes that modify survival prognosis among the recruited CF sibs (Fig. 2). Furthermore, in a case-control design the phenotypic contrast between mildly affected and severely affected sib pairs was used to detect variants that modulate CF disease severity (Table 2, Fig. 3). Finally, the pooled sample of concordant pairs was compared to the discordant siblings by case-control analysis (Fig. 4). While two distinct haplotype blocks were observed on 16p12 enabling the dissection of *SCNNIB* and *SCNNIG* mediated effects (Fig. 4), direct sequencing and haplotype breakpoint analysis was used for fine-mapping of the CF modulator on 12p13 (Table 3a, 3b and Fig. 6 within the supplement).

Skewed allele distribution among CF siblings on 12p13

Family-based analysis for association employing all CF sib pair families revealed evidence for disproportionate transmission at *TNFRSF1A*. While

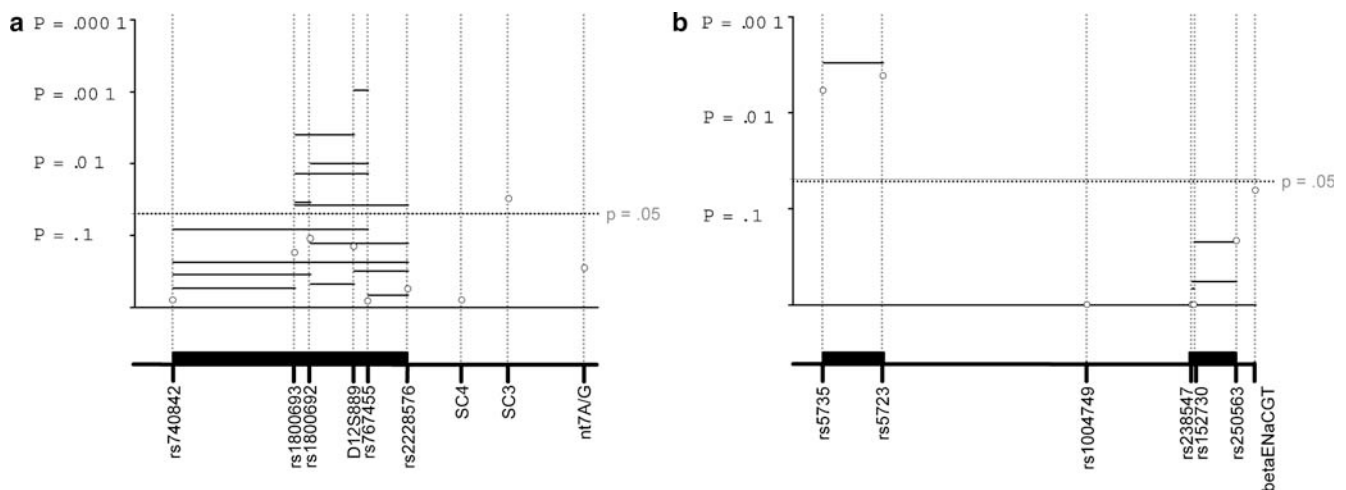


Fig. 2 Family-based analysis of haplotype blocks. Transmitted and non-transmitted alleles and haplotypes were compared for all nuclear families of 37 F508del-*CFTR* homozygous sibs pairs. **a** *TNFRSF1A* and *SCNNIA* markers on 12p13. The block *P* value for rs740842-rs1800693-rs1800692-D12S889-rs767455-rs2228576

encompassing the entire *TNFRSF1A* gene is $P=0.0085$. **b** Haplotype blocks in *SCNNIG* and *SCNNIB* on 16p12. The block *P* values are: $P=0.0043$ for rs5735-rs5723 (*SCNNIG*) and $P=0.5110$ for rs238547-rs152730-rs250563 (*SCNNIB*). The corrected *P* value for both haplotype blocks on 16p12 is $P=0.0086$

Table 2 Haplotype distribution at 12p13

| | D12S889 | rs767455 | Transmitted | Non-transmitted | Haplotype frequencies | | Dominant model | | |
|--|---------|----------|-------------|-----------------|-----------------------|-------|--------------------|--------------------|------------|
| | | | | | CON+ | CON- | CON+ | CON- | Odds ratio |
| ^a Proportion of CON+ individuals who carry at least one 10-G allele | 13 | A | 9 | 19 | 0.216 | 0.397 | | | |
| | 14 | A | 7 | 2 | 0.104 | 0.068 | | | |
| | 8 | G | 3 | 5 | 0.027 | 0.091 | | | |
| ^b Proportion of CON- individuals who carry at least one 10-G allele | 10 | G | 26 | 13 | 0.500 | 0.250 | 0.855 ^a | 0.441 ^b | 7.48 |
| | 13 | G | 0 | 7 | 0.054 | 0.080 | | | |
| | Other | | | | 0.099 | 0.115 | | | |

single-marker analysis failed to reach statistical significance after Monte Carlo simulation based correction for multiple testing ($P=0.133$), frequency distribution was significantly skewed for the haplotype block rs740842-rs1800693-rs1800692-D12S889-rs767455-rs2228576 (codon 35 of *FLJ10665* to codon 663 in exon 13 of *SCNN1A*) encompassing the entire *TNFRSF1A* gene, but only the first 35 and the last six codons of the adjacent genes *FLJ10665* and *SCNN1A*, respectively ($P=0.0085$, Fig. 2a). At D12S889-rs767455 (intron 1 to codon 12 in exon 1 of *TNFRSF1A*), haplotype (10)G was preferentially transmitted to CF offspring while haplotype (13)A was more frequently observed among non-transmitted parental haplotypes (Table 2).

Skewed allele distribution among CF siblings on 16p12

The analyzed region on 16p12 revealed disproportionate allele distributions comparing transmissions and non-transmissions (corrected $P=0.0193$). For the two SNPs within *SCNN1G*, uncorrected single locus P values peaked at rs5735 ($P=0.0059$) and rs5723 ($P=0.0041$). Two distinct haplotype blocks were detected on 16p12 (Fig. 1) and the transmission disequilibrium could be assigned to rs5735-rs5723 (codon 158 to codon 649 in *SCNN1G*) within the gene encoding the γ subunit of the epithelial sodium channel ENaC. (corrected $P=0.0086$, Fig. 2b). The rs5735-rs5723 haplotype CC was observed more frequently among CF offspring (transmitted 28 chromosomes, non-transmitted 11 chromosomes carrying the rs5735-rs5723 haplotype CC) while TG was observed more frequently among the non-transmitted parental haplotypes (transmitted 9 chromosomes, non-transmitted 21 chromosomes carrying the rs5735-rs5723 haplotype TG). Chromosomes with rs5735-rs5723 haplotypes TC and CG were observed with similar frequency on transmitted and non-transmitted chromosomes.

Association with disease severity in CF on 12p13

Association of 12p13 with CF disease severity was analyzed using case-control analysis comparing both mildly and severely affected patient pairs as well as

discordant with concordant pairs. Discordant and concordant pairs were comparable at all loci tested on 12p13 (data not shown) while mildly and severely affected pairs were highly dissimilar at 12p13. At 12p13, TGC(10)GG at rs740842-rs1800693-rs1800692-D12S889-rs767455-rs2228576 (codon 35 of *FLJ10665* to codon 663 in exon 13 of *SCNN1A*) was observed more frequently among CON+ ($P=0.02287$; Fig. 3; dominant model). Markers composing this haplotype include T35A in *FLJ10665* (rs740842) and T663A in *SCNN1A* (rs2228576), two non-synonymous nucleotide exchanges both of which failed to demonstrate association to the CF phenotype using single locus analyses (uncorrected single locus P values: 0.2568 for rs740842 and 0.6186 for rs2228576). The haplotype distributions at D12S889-rs767455 (intron 1 to codon 12 in exon 1 of *TNFRSF1A*) reflect the results of the family-based test applied to all nuclear families: the haplotype (10)G, preferentially transmitted among all nuclear families, is overrepresented among mildly affected patient pairs and the

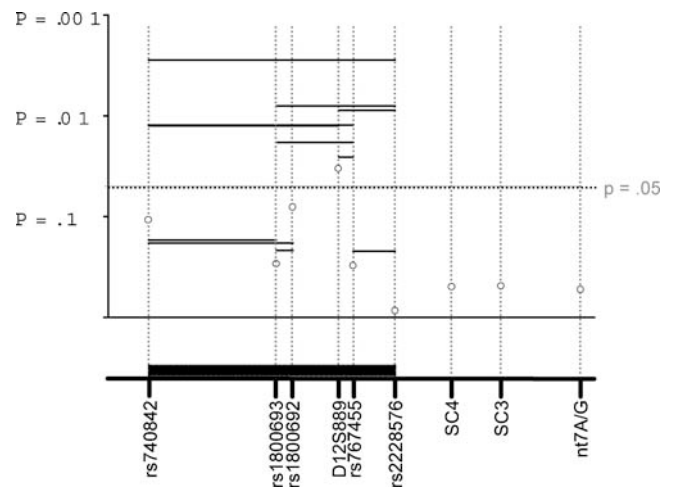


Fig. 3 Case-control analysis comparing mildly (CON+) and severely (CON-) affected patient pairs. Twelve concordant mildly affected (CON+) and 11 concordant severely affected (CON-) patient pairs were compared by case-control analysis at loci on 12p13. Data are shown for single markers (open circles) and diplotypes of adjacent markers (lines). The x-axis depicts the physical distance between the markers and the y-axis denotes the P value. $P=0.02287$ was observed for the haplotype block, as indicated by the solid black line on the physical map, of markers rs740842-rs1800693-rs1800692-D12S889-rs767455-rs2228576

haplotype (13)A, predominantly observed among non-transmitted parental haplotypes, is overrepresented in severely affected patient pairs (OR 7.48; dominant model; Table 2).

Association with intrapair discordance in CF on 16p12

Whereas mildly and severely affected patient pairs were comparable at all 16p12 loci (data not shown), case-control evaluation comparing all discordant siblings to the pooled concordant pairs showed association of *SCNN1B* markers with the CF phenotype intrapair discordance (corrected $P=0.0230$) whereby uncorrected single locus P values peaked at rs152730 in *SCNN1B* ($P=0.0031$). Haplotype block analysis revealed that the effect is mediated via rs238547-rs152730-rs250563 (codon 3 to codon 293 in *SCNN1B*) in the gene encoding the β subunit of the epithelial sodium channel ENaC (corrected $P=0.02262$, Fig. 4). rs238547-rs152730-rs250563 haplotypes were different among concordant and discordant pairs as CGC was overrepresented among discordant pairs (haplotype frequency for rs238547-rs152730-rs250563 CGC on chromosomes of concordant pairs: 0.106, on chromosomes of discordant pairs: 0.341) and haplotypes CTC and TTC were more frequent among concordant pairs (haplotype frequency for rs238547-rs152730-rs250563 CTC on chromosomes of concordant pairs: 0.232, on chromosomes of discordant pairs: 0.083; for rs238547-rs152730-rs250563 TTC on chromosomes of concordant pairs: 0.661, on chromosomes of discordant pairs: 0.483).

Sequence analysis of candidate genes *SCNN1A* and *TNFRSF1A*

The coding sequence of *SCNN1A* and flanking intron segments were analyzed by sequencing five unrelated F508del homozygotes carrying *SCNN1A* intragenic haplotypes A(10)(10)A, G(9)(10)A, G(10)(10)A, G(9)(10)G and G(10)(10)G at markers rs2228576-SC4-SC3-nt7AG (codon 663 to 5' UTR of *SCNN1A*) whereby two chromosomes for each *SCNN1A* haplotype were analyzed. A deletion in exon 13 of *SCNN1A*, F527del, was observed in one patient heterozygous for A(10)(10)A and G(10)(10)A. The other two sequenced chromosomes with these haplotypes did not carry F527del. Moreover, with the exception of T663A (rs2228576), all sequenced chromosomes displayed identical amino acid sequences throughout *SCNN1A*.

The *TNFRSF1A* sequence variant R92Q (Aksentijevich et al. 2001), reported to cause the inflammatory disorder TRAPS (tumor necrosis factor associated periodic syndrome; OMIM #142680), was previously observed on the rs1800692-rs767455 (intron 5 to codon 12 of *TNFRSF1A*) haplotype CG (Hull et al. 2002) which is part of the *TNFRSF1A* intragenic haplotype seen enriched among CON+ siblings. Analyzing the CF sibs at rs4149584, we observed the mutant 92Q in

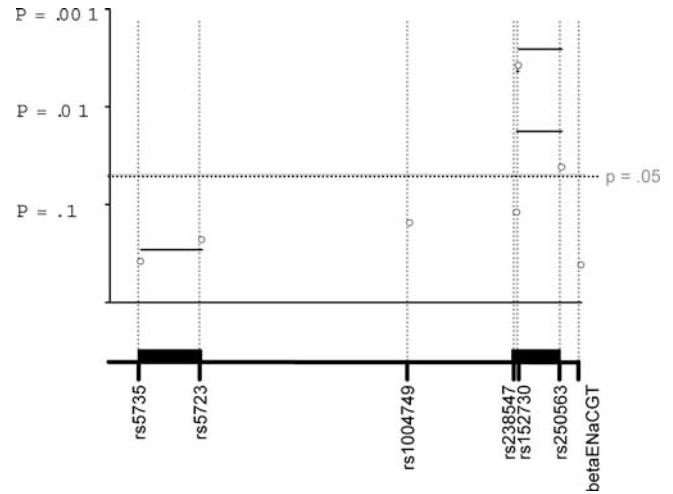


Fig. 4 Case-control analysis comparing concordant (CONC) and discordant (DIS) pairs. Twenty-three concordant patient pairs (CONC) and 14 discordant patient pairs (DIS) were compared by case-control analysis at loci on 16p12. Data are shown for single markers (open circles) or diplotypes of adjacent markers (lines). The x-axis depicts the physical distance between the markers and the y-axis denotes the P value. The two haplotype blocks are indicated by the solid lines on the physical map. $P=0.4190$ and $P=0.0086$ was observed for the haplotype blocks rs5735–rs5723 and rs238547–rs152730–rs250563, respectively. The corrected P values for the two haplotype blocks on 16p12 is $P=0.0172$

three families on chromosomes carrying the rare D12S889 alleles (8) and (9).

The genomic sequence of *TNFRSF1A* was determined by direct sequencing of four unrelated F508del homozygotes for whom the *TNFRSF1A* intragenic haplotype could be unequivocally inferred from the parental genotypes. The rs1800693-rs1800692-D12S889-rs767455 (intron 6 to codon 12 of *TNFRSF1A*) haplotype GC(10)G was represented twice in the homozygous state. Three out of the other four sequenced chromosomes carried the contrasting haplotype AT(13)A. The primary sequence of the four individuals was obtained from five partially overlapping PCR products with a total length of 12,416 bp of sequence (Table 4). Together, these fragments cover the near promoter, the 5' UTR, the entire coding sequence, the 3' UTR and the intronic sequences with the exception of 1.5 kb located centrally in intron 5. Explicitly, the entire 7.5 kb intron 1 sequence was ascertained.

The near promoter of *TNFRSF1A*, the 5' UTR and the entire *TNFRSF1A* amino acid sequence was unaltered in all sequenced samples. Apart from the already typed variants rs1800693 in intron 6 and rs1800692 in intron 5, primary sequences of GC(10)G and AT(13)A chromosomes were identical between exon 2 and the 3' UTR with the exception of a rare SNP within intron 4 observed in four families on GC(10)G intragenic background.

The sequence of the 7.5 kb intron 1 of *TNFRSF1A* differed by haplotype. Comparing the two contrasting haplotypes GC(10)G and AT(13)A, different alleles at

Table 3a Sharing of chromosomes with rs1800692-D12S889 haplotype C-10

| Phenotype | C-10 not shared ^a | C-10 shared ^b | Total number of C-10 chromosomes |
|-----------|------------------------------|--------------------------|----------------------------------|
| CON- | 7 | 1 | 9 |
| CON+ | 2 | 11 | 24 |
| DIS | 8 | 7 | 22 |
| all: | | | 55 |

^aThe parental rs1800692-D12S889 haplotype C-10 was transmitted to one sib only

^bThe parental rs1800692-D12S889 haplotype C-10 was transmitted to both sibs

seven SNPs located within intron 1 were observed. The respective seven-marker *TNFRSF1A* intron 1 haplotypes rs2284344-rs887477-rs1860545-rs4149581-rs4149580-rs4149577-D12S889-rs4149576 (all in intron of *TNFRSF1A*) were GGTGTT(10)A, observed on the four chromosomes carrying GC(10)G at rs1800693-rs1800692-D12S889-rs767455, and CTCACC(13)G, observed on the three chromosomes carrying AT(13)A at rs1800693-rs1800692-D12S889-rs767455.

Evaluation of haplotype sharing on 12p13

As outlined above, a transmission disequilibrium at *TNFRSF1A* markers was observed among all CF families and the same *TNFRSF1A* intragenic background was enriched among our concordant mildly affected siblings. We have evaluated the 12p13 haplotypes of our CF patients to describe the chromosomal segment that is shared among all CF sibs carrying the allele (10) at D12S889. Results of the haplotype reconstruction for D12S889 and the SNPs typed in *CD9*, *FLJ10665*, *TNFRSF1A* and *SCNN1A* are displayed in Tables 3a and 3b. At rs740842-D12S889 (codon 35 of *FLJ10665* to intron 1 of *TNFRSF1A*), most chromosomes carry T(10) but the rare haplotype C(10) is observed in four families indicating an ancestral recombination breakpoint between rs740842 and D12S889. For D12S889-rs2228576 (intron 1 of *TNFRSF1A* to codon 663 of *SCNN1A*), most chromosomes carry the haplotype (10)G while the recombined haplotype (10)A was seen in two families. Hence, direct reconstruction of haplotypes allowed us to

delineate the segment shared by all chromosomes carrying allele (10) at D12S889 between rs740842 in *FLJ10665* and rs2228576 in exon 13 of *SCNN1A*. Prediction of haplotypes for families with non-informative phase narrowed the shared area to a 10 kb segment defined by markers rs1800692 and D12S889, located within intron 5 and intron 1 of *TNFRSF1A*, respectively. In our CF families, all chromosomes carrying allele (10) at D12S889 share the haplotype C(10) at rs1800692-D12S889 (intron 5 to intron 1 of *TNFRSF1A*). Most chromosomes carry G(10) at rs1800693-D12S889 (intron 6 to intron 1 of *TNFRSF1A*), but the rare haplotype A(10) is predicted for two families. Toward *SCNN1A*, haplotype (10)G is observed on most chromosomes for D12S889-rs767455 (intron 1 to codon 12 of *TNFRSF1A*) while the rare haplotype (10)A was predicted for one family. As the synonymous variant rs767455 is located within exon 1 of *TNFRSF1A*, the common genomic fragment on chromosomes with allele (10) at D12S889 includes the entire intron 1 of *TNFRSF1A*.

Discussion

The amiloride-sensitive sodium channel ENaC is known for its role in body electrolyte and water regulation (Rossier et al. 2002). Mutations within ENaC subunits cause hereditary hypertension known as Liddle's syndrome (OMIM 177200) and the salt-wasting disease pseudohypoaldosteronism type 1 (OMIM 264350). Altered sodium absorption was observed in lung diseases such as respiratory distress syndrome (Barker et al 1997) and pulmonary edema (Scherrer et al. 1999). ENaC regulates the liquid balance at birth during the transition from the liquid-filled fetal lung to the air-filled organ (Barker et al. 1998) whereby all ENaC subunits are regulated at the transcriptional level (Talbot et al. 1999). We have analyzed 37 F508del-*CFTR* homozygous CF sibling pairs at 12p13 and 16p12 markers to learn whether CF disease severity is modulated by *SCNN1A*, *SCNN1B* and *SCNN1G*, encoding the three subunits of ENaC, respectively. The CF sib pairs enrolled within this study have been selected from a set of more than 300 patient pairs on the basis of their homogeneous mutation genotype in the disease-causing gene *CFTR* and their informative

Table 3b Observed alleles at 7 SNP loci on 55 chromosomes carrying allele 10 at D12S889

| Marker | rs3181301 | rs740842 | rs1800693 | rs1800692 | rs767455 | rs2228576 | nt 7A/G |
|-------------------------------|-----------------|-------------------|-----------------------|-----------------------|-------------------|------------------|------------------|
| Gene/pos. within gene | CD9/ Intron7 | FLJ10665/ T35A | TNFRSF1A/ Intron 6 | TNFRSF1A/ Intron 5 | TNFRSF1A/ P12P | SCNN1A/ T663A | SCNN1A/ 5'UTR |
| Frequent allele | A | T | G | C | G | G | A |
| Frequency on C-10 chromosomes | 0.76 | 0.80 | 0.94 | 1.00 | 0.96 | 0.94 | 0.51 |
| Rare allele | C | C | A | T | A | A | G |
| Frequency on C-10 chromosomes | 0.24 | 0.20 | 0.06 | Not observed | 0.04 | 0.06 | 0.49 |

contrasting clinical phenotype which has been classified as concordant mildly affected or concordant severely affected or discordant (Mekus et al. 2000).

Modulation of CF disease severity at 16p12

Case-control comparison revealed a modulator for CF intrapair discordance at 16p12. The haplotype distribution at rs238547-rs152730-rs250563 in *SCNN1B* was different among discordant siblings in comparison to concordant pairs (Fig. 4). Two sibs of a discordant pair mostly shared their 16p12 alleles while their phenotype is dissimilar by designation (Mekus et al. 2000). Hence, discordance must be partially mediated by interacting factors encoded elsewhere than 16p12. The value of discordant sib pairs to recognize the impact of modulators encoded beyond the loci that are in linkage disequilibrium to the investigated markers has recently been demonstrated for a CF modifier on 7q31.3 near the 3' end of the *CFTR* gene (Mekus et al. 2003). Gene-gene interactions that manifest as allelic association with the discordant phenotype in the absence of reduced intrapair sharing can be envisaged at the transcriptional level for an interaction of a regulator with the specific DNA or RNA binding site or at the post-transcriptional level through interacting proteins. Although we cannot exclude with certainty that mildly and severely affected siblings of discordant pairs carry unknown different functional variants at *SCNN1B*, both alleles being inherited on the same haplotype at 16p12, we propose that an interaction of factors encoded elsewhere in the genome with a responsive element in or near *SCNN1B* is most likely to explain the observed genotype-phenotype association at 16p12. The interaction of functionally non-equivalent variants of regulators with specific transcription factor binding sites encoded on the haplotype seen among discordant sibs would render sib pairs susceptible to discordant expression of *SCNN1B*. In this context, it is noteworthy that both reduced as well as elevated expression levels of murine *scnn1b* trigger pulmonary disease in mouse models, emphasizing the importance of *SCNN1B* expression control to avoid pathological conditions in the lung. Reduced expression of *scnn1b* renders mice susceptible to conditions analogous to pseudohypoaldosteronism type I (Pradervand et al 1999), a disease associated with excess airway liquid in humans (Kerem et al. 1999). Increased expression of *scnn1b* in airway tissue causes symptoms typically observed in CF lung disease (Mall et al. 2004). Furthermore, Thomas et al. (2002a) described an index case of systemic pseudohypoaldosteronism caused by a 1,300 bp deletion in the *SCNN1B* promoter (Thomas et al. 2004). Finally, in a murine infection model with *P. aeruginosa*, the major pathogen of chronic lung infection in CF, *scnn1b* mRNA was transiently increased (Dagenais et al. 2005) suggesting an involvement of *SCNN1B* transcriptional regulation during lung liquid clearance in host defence. In conclusion, the above observations imply

that discordance among CF sib pairs is modulated by mechanisms that control *SCNN1B* expression.

Additionally, the region analyzed on 16p12 suggests variants near or within *SCNN1G* as CF modifier. Comparing transmitted and non-transmitted alleles among all CF families (Fig. 2b), the *SCNN1G* haplotype rs5735-rs5723 TG was underrepresented among CF siblings. This contrast was not reflected by case-control comparison of concordant mildly and concordant severely affected patient pairs as TG chromosomes were rare among mildly as well as among severely affected CF patients. Assuming random transmission of parental 16p12 alleles to CF offspring, we have to interpret our failure to recruit a significant number of CF sibs carrying the rs5735-rs5723 haplotype TG as an indication of a high-risk variant on these chromosomes. In other words, our data suggest that a significant proportion of CF offspring who randomly inherited the *SCNN1G* TG haplotype from their parents did not survive to an age that allowed them to match the recruitment criteria for this study. We have only recruited pairs of which both siblings were alive. Moreover, a pulmonary function test that can be reliably conducted for children of 7 years and older was obligatory for both sibs to allow the assessment of the CF disease severity. Respiratory disease during early childhood has been related to ENaC function in infants with respiratory distress syndrome by functional analysis showing decreased sodium ion absorption (Barker et al. 1997) and expression analysis revealing decreased mRNA levels for all ENaC subunits whereby the effect was strongest for *SCNN1G* mRNA (Helve et al. 2004). Clinical data for this cross-sectional study were obtained in 1995–1996, most of the siblings being born within or prior to the early 1980s. Thus, the patients enrolled for the study have survived a decade during which CF therapy was substantially improved and mortality among CF patients decreased considerably (Corey and Farewell 1996; Goss and Rosenfeld 2004; Liou et al. 2001). It remains to be investigated in a CF population born later than our CF twins and sibs whether the improvement of the therapeutic regime could compensate for the inherent risk transmitted by 16p12 alleles mediated by *SCNN1G* and/or *SCNN1B* or by a common locus control element of these two ENaC subunits as proposed by Thomas et al. (2002b).

Modulation of CF disease severity at 12p13

While the candidate gene on 12p13 targeted within this study is *SCNN1A*, encoding the α -subunit of the ENaC, several lines of evidence indicate that the CF modulator on 12p13 is the neighboring gene *TNFRSF1A*. Explicitly, the family-based test for association enrolling all CF families (Fig. 2a) as well as the case-control comparison between concordant mildly and severely affected sib pairs (Fig. 3) identified *TNFRSF1A* intragenic markers associated with CF disease severity, but all intragenic *SCNN1A* markers failed to reach significance in these

analyses. Transmission of parental genotypes to CF siblings was markedly skewed towards the preferentially transmitted haplotype (10)G at D12S889-rs767455 (Table 2), indicating that these chromosomes carry a benign variant. In accordance with this hypothesis, the presumed benign haplotype (10)G was also observed more frequently among concordant mildly affected sib pairs (Table 2). Moreover, reconstruction of haplotypes on 12p13 denotes a shared segment restricted to *TNFRSF1A* intron 1 and flanking sequences on all chromosomes carrying allele 10 at D12S889 (Table 3b). The identification of *TNFRSF1A* as a modulator for CF disease severity is a finding by serendipity: we have started this study in 1997 by analyzing D12S889, a marker selected based on the work of Baens et al. (1995) who had mapped D12S889 and *SCNN1A* to the same cosmid clone. Towards the turn of the century, the human genome project revealed the primary sequence of the PAC clone RPC11-96H9 showing that D12S889 is actually localized within intron 1 of *TNFRSF1A*. *SCNN1A* and *TNFRSF1A* turned out to be separated by only 4 kb of intergenic sequence and consequently, we extended our candidate gene analysis at 12p13 to include this receptor for the pleiotropic cytokine TNF α that plays a major role in the regulation of immune function and inflammatory response (Wajant et al. 2003).

We sought to determine the molecular etiology underlying the modulation of CF disease severity by sequencing *TNFRSF1A* on the two frequent contrasting haplotypes (10)G and (13)A. While the coding sequence and the splice site consensus motifs of both alleles were identical, these chromosomes differed at seven SNPs within the 7.5 kb intron 1 of *TNFRSF1A*, all of which map to the segment shared on all chromosomes carrying allele 10 at D12S889 among the CF siblings (Table 3b). In this context, the findings of Dieude et al. (2004) are of interest who analyzed rs767455 in rheumatoid arthritis, a chronic inflammatory disease for which linkage to markers on 12p13 near *TNFRSF1A* have been previously reported from independent studies (Cornelis et al. 1998; MacKay et al. 2002; Jawaheer et al. 2001). Dieude et al. (2004) report that homozygosity for the allele A at rs767455 is negatively associated with rheumatoid arthritis, while the molecular mechanism could not be specified. We conclude that either one of the intron 1 SNPs or a haplotype composed of several allelic variants within intron 1 alters the functionality of the *TNFRSF1A* gene in a causative manner. Functional consequences of the *TNFRSF1A* non-coding variants can be envisaged at the DNA level because structural properties—such as bendability of the DNA helix near DNase hypersensitive sites (Brukner et al. 1995a, b) or nucleosome positioning signals (Fitzgerald et al. 1994; Baldi et al. 1996)—are determined by the primary sequence. Moreover, specific sites for DNA-binding proteins might be altered by SNPs. As a consequence, the two intron 1 variants might differ with respect to their accessibility of and responsiveness to transcriptional regulators. Alternatively or additionally, causative variants might act through the pre-mRNA as

splicing requires an interaction with RNA-binding proteins such as hnPNPs and furthermore, intron excision demands a specific tertiary structure of the pre-mRNA itself. All these mechanisms ultimately interfere with the transcriptional activation of *TNFRSF1A*, *TNFRSF1A* expression and finally with the timing and duration of TNF α signaling.

To substantiate our hypothesis that the two contrasting *TNFRSF1A* intron 1 haplotypes are not functionally equivalent we have analyzed *TNFRSF1A* for conserved non-coding sequences, presumably corresponding to regulatory elements, using the AltaVistaGenomeBrowser and applied the GenomeAtlas software to both primary sequences to identify alterations in predicted DNase hypersensitive sites and other functional properties that can be predicted from the DNA sequence (Hallin and Ussery 2004). Shifts of predicted DNaseI hypersensitivity, inverted repeats or variation within a conserved non-coding sequence were visualized for four adjacent polymorphisms out of the seven SNPs that distinguish frequent haplotypes on chromosomes with allele (10) at D12S889 and allele (13) at D12S889 (Fig. 5). Altered patterns for predicted DNaseI hypersensitive sites were observed near rs1860545-rs4149581-rs4149580 and rs887477 was located within a conserved non-coding sequence.

Implications for pathophysiology and clinic of CF

This association study identified the amiloride sensitive sodium channel ENaC and the TNF α receptor TNFR1 as modifiers in CF. Transmission disequilibrium of haplotype blocks was observed for the whole study population at both investigated loci implying that carriers of risk haplotypes were underrepresented when we recruited in 1995–1996 F508del-*CFTR* homozygous twins and siblings of at least 7 years of age. Median survival has increased from less than 10 years in the late 1970s to more than 30 years in the present patient population (Corey and Farewell 1996; Goss and Rosenfeld 2004) and correspondingly we can expect larger proportions of risk alleles in today's younger birth cohorts if patients from schoolchildren age to early adulthood are investigated. The depletion of risk alleles that was particularly prominent in the mildly affected subgroup of our study cohort is a strong argument that ENaC and TNFR1 are clinically highly relevant genetic modulators of CF disease.

Absent or dysfunctional CFTR manifests in impaired apical chloride conductance as the hallmark of the basic defect in CF (reviewed by Vankeerberghen et al. 2002). CFTR and ENaC are tightly linked in the regulation of ion and water homeostasis in exocrine epithelia, particularly in the antagonistic control of the water content of the airway surface liquid (ASL, reviewed by Boucher 2004). The steady state is likely maintained by a balance between sodium absorption and chloride secretion. Correspondingly both the overexpression of ENaC or downregulation of CFTR lead to ASL volume

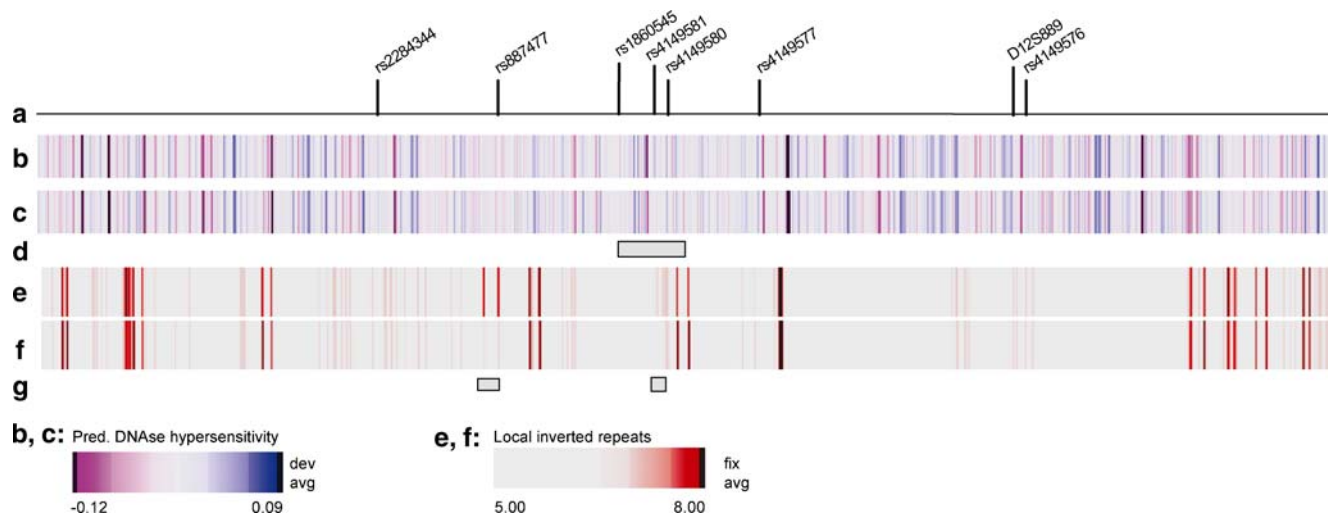


Fig. 5 Functional annotation of *TNFRSF1A* intron 1. The illustration shows the entire 7.5 kb of *TNFRSF1A* with allocated genetic markers (a). Putative functional elements were visualized using the GenomeAtlas software (b, c, e, f). Primary sequences for 8-marker haplotypes at rs2284344-rs887477-rs1860545-rs4149581-rs4149580-rs4149577-D12S889-rs4149576 were compared with respect to predicted DNase hypersensitive sites based on the trinucleotide model of Brukner et al. (1995a, 1995b). Results for haplotype GGTGTT(10)A are shown in (b) and for

CTCACC(13)G are shown in (c). Comparing the two haplotypes, patterns for predicted DNase hypersensitive sites were different near rs1860545, rs4149581 and rs4149580 as indicated by gray boxes in (d). Primary sequences were monitored for local inverted repeats in (e) for haplotype GGTGTT(10)A and (f) for haplotype CTCACC(13)G, whereby differences between the two haplotypes are indicated by gray boxes in (g) near rs887477 and rs4149581-rs4149580

depletion. The genetic data of our association study support current concepts of the molecular and cellular pathophysiology of CF lung disease (Boucher 2004) that the ENaC is a major modulator of the basic defect in CF.

Depletion of a specific compartment of the ASL, i.e. the periciliary liquid, appears to abrogate both cilia-dependent and cough clearance in CF and leads to abnormal mucus transport which predispose the lung to infection with opportunistic bacterial pathogens, the typical secondary complication of CF lung disease. TNF α thereby plays a pivotal role in orchestrating innate inflammatory responses towards the opportunistic bacterial pathogens that colonize the CF airways. Upon binding of TNF α , TNFR1 signaling triggers several intracellular signaling pathways which control gene expression through transcription factors such as NF- κ B and AP-1 (reviewed by Wajant et al. 2003). This pleiotropic action leads to local expression of chemokines and cytokines, promoting the adhesion, extravasation, attraction and activation of leukocytes at the site of infection. Later, TNF α facilitates transition from innate to acquired immunity by enhancing antigen presentation and T cell costimulation (Wajant et al. 2003). While the cardinal role of TNF α is to stimulate inflammation, the cytokine is also capable to induce apoptosis through TNFR1 when transcription, translation or NF- κ B signaling are blocked (reviewed by Varfolomeev and Ashkenazi 2004). In summary, both the proinflammatory and the proapoptotic TNF α signal is transmitted via TNFR1. These actions are beneficial for the healthy immunocompetent host, but in case of CF they reinforce chronic airway inflammation: The bacterial pathogens evade the

host response and sustain a vicious cycle of fooled self-destructive host defense characterized by a pronounced imbalance between pro- and anti-inflammatory cytokines (reviewed in Tümmler and Kiewitz 1999).

Moreover, homozygosity for F508del-*CFTR* is an endogenous proinflammatory condition. Most newly synthesized F508del-*CFTR* cannot pass the ER quality control because of improper folding and becomes ubiquitinated and degraded in the proteasome (Vankeerberghen et al. 2002). The congestion of the ER with mutant F508del-*CFTR* may cause the proinflammatory ER overload response that ultimately results in the activation of NF- κ B (Knorre et al. 2002). Thus, the transcription factor NF- κ B that plays an important integrating role in the intracellular regulation of immune response and inflammation, is induced in F508del-*CFTR* homozygotes by two major stimuli, the mutant *CFTR* protein and the TNF α signaling cascade.

In summary, the body of molecular and cellular pathophysiological data provides plain evidence for key roles of the TNF α receptor TNFR1 and the epithelial sodium channel ENaC for the pathogenesis of CF disease that is now substantiated by the genetic data gained in this association study. Sequence variants of the ENaC and TNFR1 genes are associated with the severity of CF and most likely modify basic defect and inflammation in CF patients. Since this study identified *SCNN1B*, *SCNN1G* and *TNFRSF1A* as clinically relevant genetic modifiers, they are candidates for therapeutic targets. ENaC and TNFR1 are components of the network that sustains the vicious cycle of infection and inflammation in CF airways and thereby determines morbidity and

mortality in most individuals with CF. Clinical investigations in carriers with the beneficial haplotype blocks may help to define the surrogate parameters of how ENaC or TNFR1 should be modulated to gain a beneficial therapeutic effect.

Electronic database resources

The cystic fibrosis mutation database

available at: <http://www.gent.sickkids.on.ca/cftr/>

Genomic sequence and annotation of NT_009756

(International Human Genome Sequencing Consortium 2003)

and

Genomic sequence and annotation of NT_010393

(International Human Genome Sequencing Consortium 2004)

is the same:

available at: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide>

The Vista Genome Browser

available at: <http://www.pipeline.lbl.gov/cgi-bin/gateway2>

Acknowledgements This work was executed as part of the European Cystic Fibrosis Twin and Sibling Study and financially supported by the BIOMED programme of the European Union (QL G1-CF-2001-01005), the Deutsche Forschungsgemeinschaft (SFB621-C7), the Deutsche Fördergesellschaft für die Mukoviszidoseforschung eV and the Mukoviszidose eV. TB was supported by grant KN 378/1-1 (Project D1 of FOR 423) from the Deutsche Forschungsgemeinschaft. The authors are indebted to the patients and families for their participation in the study and gratefully acknowledge the co-operation and assistance by the physicians, CF centers and their staff without which this study would not have been possible.

References

- Aksentijevich I, Galon J, Soares M, Mansfield E, Hull K, Oh HH, Goldbach-Mansky R, Dean J, Athreya B, Reginato AJ, Henrickson M, Pons-Estel B, O'Shea JJ, Kastner DL (2001) The tumor-necrosis-factor receptor-associated periodic syndrome: new mutations in TNFRSF1A, ancestral origins, genotype-phenotype studies, and evidence for further genetic heterogeneity of periodic fevers. *Am J Hum Genet* 69:301–314
- Baens M, Aerssens J, van Zand K, Van den Berghe H, Marynen P (1995) Isolation and regional assignment of human chromosome 12p cDNAs. *Genomics* 29:44–52
- Baldi P, Brunak S, Chauvin Y, Krogh A (1996). Naturally occurring nucleosome positioning signals in human exons and introns. *J Mol Biol* 263:503–510
- Barker PM, Gowen CW, Lawson EE, Knowles MR (1997) Decreased sodium ion absorption across nasal epithelium of very premature infants with respiratory distress syndrome. *J Pediatr* 130:373–377
- Barker PM, Nguyen MS, Gatzky JT, Grubb B, Norman H, Hummler E, Rossier B, Boucher RC, Koller B (1998) Role of gammaENaC subunit in lung liquid clearance and electrolyte balance in newborn mice Insights into perinatal adaptation and pseudohypoaldosteronism. *J Clin Invest* 102:1634–1640
- Becker T, Knapp M (2004a) Maximum-likelihood estimation of haplotype frequencies in nuclear families. *Genet Epidemiol* 27:21–32
- Becker T, Knapp M (2004b) A powerful strategy to account for multiple testing in the context of haplotype analysis. *Am J Hum Genet* 75:561–570
- Becker T, Cichon S, Jönson E, Knapp M (2005) Multiple testing in the context of haplotype analysis revisited: application to case-control data. *Ann Hum Genet* 69:747–756
- Boucher RC (2004) New concepts of the pathogenesis of cystic fibrosis lung disease. *Eur Respir J* 23:146–158
- Brukner I, Sanchez R, Suck D, Pongor S (1995a) Sequence-dependent bending propensity of DNA as revealed by DNase I: parameters for trinucleotides. *EMBO J* 14:1812–1818
- Brukner I, Sanchez R, Suck D, Pongor S (1995b) Trinucleotide models for DNA bending propensity: comparison of models based on DNase I digestion and nucleosome packaging data. *J Biomol Struct Dyn* 132:309–317
- Chang SS, Grunder S, Hanukoglu A, Rosler A, Mathew PM, Hanukoglu I, Schild L, Lu Y, Shimkets RA, Nelson-Williams C, Rossier BC, Lifton RP (1996) Mutations in subunits of the epithelial sodium channel cause salt wasting with hyperkalaemic acidosis, pseudohypoaldosteronism type 1. *Nat Genet* 12:248–253
- Corey M, Farewell V (1996) Determinants of mortality from cystic fibrosis in Canada, 1970–1989. *Am J Epidemiol* 143:1007–1017
- Cornelis F, Faure S, Martinez M, Prud'homme JF, Fritz P, Dib C, Alves H, Barrera P, de Vries N, Balsa A, Pascual-Salcedo D, Maenaut K, Westhovens R, Migliorini P, Tran TH, Delaye A, Prince N, Lefevre C, Thomas G, Poirier M, Soubigou S, Alibert O, Lasbleiz S, Fouix S, Bouchier C, Liote F, Loste MN, Lepage V, Charron D, Gyapay G, Lopes-Vaz A, Kuntz D, Bardin T, Weissenbach J, ECRAF (1998) New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. *Proc Natl Acad Sci USA* 95:10746–10750
- Cutting GR (2005) Modifier genetics: cystic fibrosis. *Annu Rev Genomics Hum Genet* 6:237–260
- Dagenais A, Gosselin D, Guilbault C, Radzioch D, Berthiaume Y (2005) Modulation of epithelial sodium channel (ENaC) expression in mouse lung infected with *Pseudomonas aeruginosa*. *Respir Res* 6:2
- Dieude P, Osorio J, Petit-Teixeira E, Moreno S, Garnier S, Cailleau-Moindrault S, Stalens C, Lasbleiz S, Bardin T, Prum B, Cornelis F; European Consortium on Rheumatoid Arthritis Families (2004) A TNFR1 genotype with a protective role in familial rheumatoid arthritis. *Arthritis Rheum* 50:413–419
- Dolan CV, Boomsma DI (1998) Optimal selection of sib pairs from random samples for linkage analysis of a QTL using the EDAC test. *Behav Genet* 28:197–206
- Eaves L, Meyer J (1994) Locating human quantitative trait loci: guidelines for the selection of sibling pairs for genotyping. *Behav Genet* 24:443–455
- Eidelman O, Srivastava M, Zhang J, Leighton X, Murtie J, Jozwik C, Jacobson K, Weinstein DL, Metcalf EL, Pollard HB (2001) Control of the proinflammatory state in cystic fibrosis lung epithelial cells by genes from the TNF- α /NF- κ B pathway. *Mol Med* 7:523–534
- European Working Group on Cystic Fibrosis Genetics (1990) Gradient of distribution in Europe of the major CF mutation and of its associated haplotype. *Hum Genet* 85:436–445
- Fitzgerald DJ, Dryden GL, Bronson EC, Williams JS, Anderson JN (1994) Conserved patterns of bending in satellite and nucleosome positioning DNA. *J Biol Chem* 269:21303–21314
- Fuchs P, Strehl S, Dworzak M, Himmler A, Ambros PF (1992) Structure of the human TNF receptor 1 (p60) gene (TNFR1) and localization to chromosome 12p13. *Genomics* 13:219–224
- Goss CH, Rosenfeld M (2004) Update on cystic fibrosis epidemiology. *Curr Opin Pulm Med* 10:510–514
- Gross-Bellard M, Dudet P, Chambon P (1973) Isolation of high molecular weight DNA from mammalian cells. *Eur J Biochem* 36:32–38
- Hallin PF, Ussery DW (2004) CBS Genome Atlas Database: a dynamic storage for bioinformatic results and sequence data. *Bioinformatics* 20:3682–3686
- Hanukoglu A, Bistrizter T, Rakover Y, Mandelberg A (1994) Pseudohypoaldosteronism with increased sweat and saliva electrolyte values and frequent lower respiratory tract infections mimicking cystic fibrosis. *J Pediatr* 125:752–755

- Helve O, Pitkanen OM, Andersson S, O'Brodovich H, Kirjavainen T, Otulakowski G (2004) Low expression of human epithelial sodium channel in airway epithelium of preterm infants with respiratory distress. *Pediatrics* 113:1267–1272
- Hull KM, Drewe E, Akseptijevich I, Singh HK, Wong K, McDermott EM, Dean J, Powell RJ, Kastner DL (2002) The TNF receptor-associated periodic syndrome (TRAPS): emerging concepts of an autoinflammatory disorder. *Medicine* 81:349–368
- International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431:931–945
- Jawaheer D, Seldin MF, Amos CI, Chen WV, Shigeta R, Monteiro J, Kern M, Criswell LA, Albani S, Nelson JL, Clegg DO, Pope R, Schroeder HW Jr, Bridges SL Jr, Pisetsky DS, Ward R, Kastner DL, Wilder RL, Pincus T, Callahan LF, Flemming D, Wener MH, Gregersen PK (2001) A genomewide screen in multiplex rheumatoid arthritis families suggests genetic overlap with other autoimmune diseases. *Am J Hum Genet* 68:927–936
- Ji HL, Chalfant ML, Jovov B, Lockhart JP, Parker SB, Fuller CM, Stanton BA, Benos DJ (2000) The cytosolic termini of the beta- and gamma-ENaC subunits are involved in the functional interactions between cystic fibrosis transmembrane conductance regulator and epithelial sodium channel. *J Biol Chem* 275:27947–27956
- Johannsen HK, Nir M, Hoiby N, Koch C, Schwartz M (1991) Severity of cystic fibrosis in patients homozygous and heterozygous for the deltaF508 mutation. *Lancet* 337:631–634
- Kerem E, Corey M, Kerem BS, Rommens J, Markiewicz D, Levison H, Tsui LC, Durie P (1990) The relationship between genotype and phenotype in cystic fibrosis – analysis of the most common mutation (deltaF508). *N Engl J Med* 323:1517–1522
- Kerem E, Bistrizter T, Hanukoglu A, Hofmann T, Zhou Z, Bennett W, MacLaughlin E, Barker P, Nash M, Quittell L, Boucher R, Knowles MR (1999) Pulmonary epithelial sodium-channel dysfunction and excess airway liquid in pseudohypoaldosteronism. *N Engl J Med* 341:156–162
- Knapp M, Becker T (2003) Family-based association analysis with tightly linked markers. *Hum Hered* 56:2–9
- Knorre A, Wagner M, Schaefer HE, Colledge WH, Pahl HL (2002) DeltaF508-CFTR causes constitutive NF- κ B activation through an ER-overload response in cystic fibrosis lungs. *Biol Chem* 383:271–282
- Lewontin RC (1964) The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* 49:49–67
- Liou TG, Adler FR, FitzSimmons SC, Cahill BC, Hibbs JR, Marshall BC (2001) Prediction of 5-year survivorship model of cystic fibrosis. *Am J Epidemiol* 153:345–352
- Littlewood JM (2000) Good care for people with cystic fibrosis. *Paediatr Respir Rev* 1:179–89
- MacKay K, Eyre S, Myerscough A, Milicic A, Barton A, Laval S, Barrett J, Lee D, White S, John S, Brown MA, Bell J, Silman A, Ollier W, Wordsworth P, Worthington J (2002) Whole-genome linkage analysis of rheumatoid arthritis susceptibility loci in 252 affected sibling pairs in the United Kingdom. *Arthritis Rheum* 46:632–639
- Mall M, Grubb BR, Harkema JR, O'Neal WK, Boucher RC (2004) Increased airway epithelial Na⁺ absorption produces cystic fibrosis-like lung disease in mice. *Nat Med* 10:487–493
- Marthinsen L, Kornfalt R, Aili M, Andersson D, Westgren U, Schaedel C (1998) Recurrent *Pseudomonas bronchopneumonia* and other symptoms as in cystic fibrosis in a child with type I pseudohypoaldosteronism. *Acta Paediatr* 87:472–474
- Mekus F, Dörk T, Deufel T, Morral N, Tümmler B (1995) Analysis of microsatellites by direct blotting electrophoresis and chemiluminescence detection. *Electrophoresis* 16:1886–1888
- Mekus F, Ballmann M, Bronsveld I, Bijman J, Veeze H, Tümmler B (2000) Categories of deltaF508 homozygous cystic fibrosis twin and sibling pairs with distinct phenotypic characteristics. *Twin Res* 3:277–293
- Mekus F, Laabs U, Veeze H, Tümmler B (2003) Genes in the vicinity of *CFTR* modulate the cystic fibrosis phenotype in highly concordant or discordant F508del homozygous sib pairs. *Hum Genet* 112:1–11
- Pradervand S, Barker PM, Wang Q, Ernst SA, Beerhoff F, Grubb BR, Burnier M, Schmidt A, Bindels RJ, Gatzky JT, Rossier BC, Hummler E (1999) Salt restriction induces pseudohypoaldosteronism type 1 in mice expressing low levels of the beta-subunit of the amiloride-sensitive epithelial sodium channel. *Proc Natl Acad Sci USA* 96:1732–1737
- Risch N, Zhang H (1995) Extreme discordant sib pairs for mapping quantitative trait loci in humans. *Science* 268:1584–1589
- Rommens J, Kerem BS, Greer W, Chang P, Tsui LC, Ray P (1990) Rapid nonradioactive detection of the major cystic fibrosis mutation. *Am J Hum Genet* 46:395–396
- Rossier BC, Pradervand S, Schild L, Hummler E (2002) Epithelial sodium channel and the control of sodium balance: interaction between genetic and environmental factors. *Annu Rev Physiol* 64:877–897
- Santis G, Osborne L, Knight RA, Hodson ME (1990) Independent genetic determinants of pancreatic and pulmonary status in cystic fibrosis. *Lancet* 336:1081–1084
- Scherrer U, Sartori C, Lepori M, Allemann Y, Duplain H, Trueb L, Nicod P (1999) High-altitude pulmonary edema: from exaggerated pulmonary hypertension to a defect in transepithelial sodium transport. *Adv Exp Med Biol* 474:93–107
- Shimkets RA, Warnock DG, Bosisit CM, Nelson-Williams C, Hansson JH, Schambelan M, Gill JR Jr, Ulick S, Milora RV, Findling JW, Canessa CM, Rossier BC, Lifton RP (1994) Liddle's syndrome: heritable human hypertension caused by mutations in the beta subunit of the epithelial sodium channel. *Cell* 79:407–414
- Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52:506–516
- Strautnieks SS, Thompson RJ, Hanukoglu A, Dillon MJ, Hanukoglu I, Kuhnle U, Seckl J, Gardiner RM, Chung E (1996) Localisation of pseudohypoaldosteronism genes to chromosome 16p12.2-13.11 and 12p13.1-pter by homozygosity mapping. *Hum Mol Genet* 5:293–299
- Tabary O, Escotte S, Couetil JP, Hubert D, Dusser D, Puchelle E, Jacquot J (2001) Relationship between $\text{I}\kappa\text{B}\alpha$ deficiency, NF κB activity and interleukin-8 production in CF human airway epithelial cells. *Pflügers Arch* 443(Suppl 1):S40–S44
- Talbot CL, Bosworth DG, Briley EL, Fenstermacher DA, Boucher RC, Gabriel SE, Barker PM (1999) Quantitation and localization of ENaC subunit expression in fetal, newborn, and adult mouse lung. *Am J Respir Cell Mol Biol* 20:398–406
- Thomas CP, Zhou J, Liu KZ, Mick VE, MacLaughlin E, Knowles M (2002a) Systemic pseudohypoaldosteronism from deletion of the promoter region of the human Beta epithelial Na(+) channel subunit. *Am J Respir Cell Mol Biol* 27:314–319
- Thomas CP, Loftus RW, Liu KZ, Itani OA (2002b) Genomic organization of the 5' end of human beta-ENaC and preliminary characterization of its promoter. *Am J Physiol Renal Physiol* 282:F898–F909
- Thomas CP, Zhou J, Liu KZ, Mick VE, MacLaughlin E, Knowles M (2004) Gene symbol: SCNN1B. Disease: Pseudohypoaldosteronism, type 1. *Hum Genet* 114:402
- Tümmler B, Kiewitz C (1999) Cystic fibrosis: an inherited susceptibility to bacterial respiratory infections. *Mol Med Today* 5:351–358
- Vankeerberghen A, Cuppens H, Cassiman JJ (2002) The cystic fibrosis transmembrane conductance regulator: an intriguing protein with pleiotropic functions. *J Cyst Fibros* 1:13–29
- Varfolomeev EE, Ashkenazi A (2004) Tumor necrosis factor: an apoptosis JuNKie? *Cell* 116:491–497
- Voilley N, Bassilana F, Mignon C, Merscher S, Mattei MG, Carle GF, Lazdunski M, Barbry P (1995) Cloning, chromosomal localization, and physical linkage of the beta and gamma subunits (SCNN1B and SCNN1G) of the human epithelial amiloride-sensitive sodium channel. *Genomics* 28:560–565
- Wajant H, Pfizenmaier K, Scheurich P (2003) Tumor necrosis factor signaling. *Cell Death Differ* 10:45–65
- Welsh MJ, Tsui LC, Boat TF, Beaudet AL (1995) Cystic fibrosis. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) *The metabolic basis of inherited disease*. McGraw-Hill, New York, pp 3799–3876